

ANALISIS INTERVENSI DAN DETEKSI OUTLIER PADA DATA WISATAWAN DOMESTIK (Studi Kasus di Daerah Istimewa Yogyakarta)

Lenny Budiarti¹, Tarno², Budi Warsito³

¹Mahasiswa Jurusan Statistika FSM UNDIP

^{2,3}Staf Pengajar Jurusan Statistika FSM UNDIP

ABSTRAK

Data wisatawan sangat menarik untuk dilakukan kajian karena sektor kepariwisataan Indonesia merupakan penggerak perekonomian nasional yang potensial untuk memacu pertumbuhan perekonomian yang lebih tinggi di masa yang akan datang. Oleh sebab itu, ramalan tentang data wisatawan sangat diperlukan bagi pelaku bisnis dalam bidang pariwisata. Data wisatawan cenderung fluktuatif hal ini disebabkan oleh banyak faktor yang mempengaruhi jumlah kunjungan wisatawan di suatu daerah secara ekstrim, seperti bencana alam, peraturan pemerintah, kestabilan sosial, kerusakan dan terorisme. Data ekstrim seperti itu dapat dikaji menggunakan analisis intervensi dan deteksi outlier. Model intervensi adalah suatu model *time series* yang dapat digunakan untuk memodelkan dan meramalkan data yang mengandung intervensi baik dari faktor internal maupun eksternal. Di dalam model intervensi terdapat dua fungsi yaitu fungsi *step* dan *pulse*. Fungsi *step* merupakan suatu bentuk intervensi yang terjadi dalam kurun waktu yang panjang sedangkan fungsi *pulse* adalah suatu bentuk intervensi yang terjadi hanya dalam suatu waktu tertentu. Untuk deteksi outlier terdapat empat tipe outlier yaitu *additive outlier* (AO), *innovational outlier* (IO), *level shift* (LS) dan *temporary change* (TC). Sebagai studi kasus dilakukan kajian empiris terhadap data wisatawan domestik ke DI Yogyakarta periode Januari 2006 sampai dengan Desember 2010 yang menginap di hotel berbintang dan hotel melati yang ada di seluruh DI Yogyakarta. Berdasarkan hasil penelitian yang telah dilakukan, diketahui bahwa intervensi terjadi pada Januari 2010 dengan menggunakan fungsi *pulse* dengan nilai MSE sebesar 1172. Sedangkan berdasarkan intervensi dan deteksi outlier yang dilakukan diketahui terdapat lima buah outlier namun hanya empat outlier yang signifikan masuk kedalam model intervensi dengan nilai MSE sebesar 523,7167. Sehingga model intervensi dan deteksi outlier dipilih sebagai model terbaik.

Kata kunci: Wisatawan domestik, model intervensi, fungsi *pulse*, deteksi outlier

ABSTRACT

The tourist data is very interesting to be studied because the Indonesian tourism sector is an activator of the national economic which is potential to push higher economic growth in the future. Therefore, the forecast about tourist data is very needed for tourism business. The tourist data tend to fluctuate caused by many factors that affect the number of tourists extremely in an area, such as disasters, government regulation, social stability, violence and terrorism. That the extreme data can be assessed using intervention analysis and outlier detection. Intervention model is a time series model that can be used to forecast data consist of intervention of internal and external factors. In the intervention model, there are two kinds of intervention function, i.e., step and pulse functions. Step function is a form of intervention occurred in period of time while the pulse function is a form of intervention occurred only in a certain time. For the outlier detection, there are four types, such as additive outlier (AO), innovational outlier (IO), level shift (LS) and temporary change (TC). As an empirical studies was conducted by the domestic tourists data in Yogyakarta from January 2006 until December 2010 who staying on five-star hotels and motel throughout Yogyakarta. Based on the result of this research, known that the intervention occurred on January 2010 using the pulse function with MSE value 1172. Meanwhile based on the outliers detection, known any five outliers but only four outliers that significant included to the intervention model with MSE value 523,7167. So, the intervention model and outlier detection are chosen as a the best model based on the smallest MSE criterion.

Keywords: Domestic tourists, intervention model, pulse function, outlier detection

1. PENDAHULUAN

Analisis *time series* merupakan salah satu prosedur statistika yang diterapkan untuk meramalkan struktur probabilitas keadaan yang akan datang dalam rangka pengambilan keputusan. Pemodelan *time series* seringkali dikaitkan dengan proses peramalan (*forecasting*) suatu nilai karakteristik tertentu pada periode mendatang. Peramalan adalah menduga atau memperkirakan suatu keadaan di masa yang akan datang berdasarkan keadaan masa lalu dan sekarang yang diperlukan untuk menetapkan kapan suatu peristiwa akan terjadi, sehingga tindakan yang tepat dapat dilakukan (Makridakis, dkk., 1999). Pada umumnya peramalan *time series* dilakukan dengan menggunakan metode *Autoregressive Integrated Moving Average* (ARIMA), *exponential smoothing*, dekomposisi atau regresi. Meskipun pendekatan semacam ini efisien untuk peramalan *time series*, namun masih menunjukkan kekurangan ketika terjadi gangguan *noise* atau data-data yang berfluktuasi ekstrim. Data yang berfluktuasi ekstrim dapat mengindikasikan adanya suatu intervensi ataupun outlier. Fluktuasi yang ekstrim tersebut dapat disebabkan oleh berbagai faktor baik faktor eksternal maupun faktor internal, seperti bencana alam, peraturan pemerintah, kestabilan ekonomi, kerusakan, dan terorisme. Data ekstrim tersebut dapat dikaji dengan analisis intervensi dan deteksi outlier. Model intervensi adalah salah satu metode *time series* yang biasanya digunakan untuk menjelaskan pengaruh intervensi faktor eksternal maupun internal terhadap suatu data *time series* (Suhartono, 2007). Dalam deteksi outlier terdapat empat tipe outlier yaitu *additive outlier* (AO), *innovational outlier* (IO), *level shift* (LS) dan *temporary change* (TC).

Data wisatawan sangat menarik untuk dilakukan kajian karena sektor kepariwisataan Indonesia merupakan penggerak perekonomian nasional yang potensial untuk memacu pertumbuhan perekonomian yang lebih tinggi di masa yang akan datang. Ramalan akan jumlah kunjungan wisatawan sangat diperlukan bagi pelaku bisnis dalam bidang pariwisata misalnya untuk jasa perhotelan, restoran, travel dan lain-lain. Dalam penulisan ini akan dicoba untuk memodelkan dan meramalkan data kunjungan wisatawan domestik ke Yogyakarta dengan metode *time series* intervensi untuk melihat dampak penurunan pengunjung pada bulan-bulan berikutnya serta melihat besar dan lamanya efek intervensi yang terjadi. Selain itu untuk mengatasi outlier yang ada dalam data digunakan deteksi outlier. Data yang digunakan adalah data jumlah kunjungan wisatawan domestik ke Yogyakarta setiap bulannya dari tahun 2006 sampai dengan 2010 yang diambil dari Jogja dalam angka (*Jogja in Figure*) tahun 2006 sampai dengan tahun 2010 yang diterbitkan oleh Badan Pusat Statistik. Data yang diamati hanya jumlah wisatawan domestik yang menginap di hotel berbintang dan hotel melati yang ada di seluruh Yogyakarta. Pemodelan dan peramalan yang dibahas adalah model intervensi dan deteksi outlier.

2. TINJAUAN PUSTAKA

2.1 Analisis *Time series*

Analisis *time series* dikenalkan pada tahun 1970 oleh George E. P. Box dan Gwilym M. Jenkins melalui bukunya *Time series Analysis: Forecasting and Control*. Sejak saat itu, *time series* mulai banyak dikembangkan. Dasar pemikiran *time series* adalah pengamatan sekarang (Z_t) tergantung pada 1 atau beberapa pengamatan sebelumnya (Z_{t-k}). Dengan kata lain, model *time series* dibuat karena secara statistik ada korelasi (dependensi) antar deret pengamatan. Untuk melihat adanya korelasi antar pengamatan, dapat dilakukan uji korelasi antar pengamatan yang sering dikenal dengan *autocorrelation function* (ACF). Tujuan analisis *time series* antara lain memahami dan menjelaskan mekanisme tertentu, meramalkan suatu nilai di masa depan, dan mengoptimalkan sistem kendali. Analisis *time series* dapat diterapkan di bidang ekonomi, bisnis, industri, teknik dan ilmu-ilmu sosial (Makridakis *et.al*, 1998).

Berbagai metode telah dikembangkan dalam mengolah data *time series* untuk memperoleh suatu model yang memberikan hasil ramalan yang lebih akurat. Metode yang digunakan antara lain adalah metode ARIMA Box-Jenkins (Box dan Jenkins, 1976) yang digunakan untuk mengolah *time series* yang univariat dan metode analisis fungsi transfer digunakan untuk mengolah data *time series* multivariat. Di dalam pembentukan model fungsi transfer digunakan metode ARIMA Box-Jenkins untuk menggabungkan deret-deret input (Z_t) dan input-input lain yang digabungkan dalam satu kelompok yang disebut

“gangguan” atau noise (n_t). Untuk dapat diolah dengan menggunakan metode ARIMA Box-Jenkins, suatu data *time series* harus memenuhi syarat stasioneritas.

Misal Z_1, Z_2, \dots, Z_t merupakan proses stokastik untuk runtun waktu diskrit. Proses di atas disebut stasioner jika mean dan variansinya konstan untuk setiap titik t dan kovarian yang konstan untuk setiap selang waktu k

$$1. E(Z_t) = \mu \quad \text{konstan untuk semua } t$$

$$2. \text{Var}(Z_t) = \sigma^2 \quad \text{konstan untuk semua } t$$

$$3. \text{Cov}(Z_t, Z_{t+k}) = \gamma_k \quad \text{konstan untuk semua } t \text{ dan semua } k \neq 0 \quad \gamma_k \text{ adalah autokovariansi pada lag } k. \text{ (Soejoeti, 1987).}$$

Umumnya untuk mengatasi ketidakstasioneran terhadap varians dilakukan Transformasi Box-Cox yang didefinisikan sebagai berikut (Wei, 2006)

$$T(Z_t) = Z_t^{(\lambda)} = \frac{Z_t^\lambda - 1}{\lambda}$$

Sedangkan ketidakstasioneran terhadap mean dapat diatasi dengan diferensi. Uji stasioneritas terhadap mean dapat dilakukan melalui uji Unit Root atau Dickey-Fuller Test.

Salah satu model matematis yang umum digunakan dalam pemodelan time series adalah model ARIMA Box-Jenkins yang didefinisikan sebagai berikut.

$$\phi_p(B)(1-B)^d Z_t = \theta_q(B)a_t$$

dimana $\phi_p(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$ merupakan operator AR yang stasioner dan $\theta_q(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$ merupakan operator MA yang *invertible*. Orde p, d, q didapatkan dari lag yang signifikan pada plot pada plot *Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF).

2.2 Analisis Intervensi

Dalam praktek seringkali ditemukan data *time series* yang dipengaruhi kejadian-kejadian khusus. Kejadian khusus yang dimaksud di sini adalah adanya suatu intervensi baik yang bersifat eksternal maupun internal yang mempengaruhi pola data. Pada analisis intervensi, diasumsikan bahwa kejadian intervensi terjadi pada waktu T yang diketahui dari suatu *time series* (Box *et al.*, 1994). Tujuan utama dari analisis ini adalah mengukur besar dan lamanya efek intervensi pada suatu *time series* (Wei, 2006). Bentuk umum dari model intervensi adalah

$$Z_t = \theta_0 + \sum_{j=1}^k \frac{\omega_{sj}(B) B^{bj}}{\delta_{rj}(B)} I_{jt} + \frac{\theta(B)}{\phi(B)} a_t$$

dengan

Z_t : variabel respon pada saat t

θ_0 : konstanta

j : banyaknya intervensi yang terjadi, $j = 1, 2, \dots, k$

I_{jt} : variabel intervensi

b : delay waktu mulai terjadi efek intervensi

$\omega_s(B) : \omega_0 - \omega_1 B - \dots - \omega_s B^s$ (s menunjukkan lamanya suatu intervensi berpengaruh pada data setelah b periode)

$\delta_r(B) : 1 - \delta_1(B) - \dots - \delta_r B^r$ (r pola efek intervensi yang terjadi setelah $b+s$ periode sejak kejadian intervensi pada waktu T)

$\frac{\theta(B)}{\phi(B)} a_t$: noise yang berupa model ARIMA tanpa adanya pengaruh intervensi

Secara umum ada dua macam variabel intervensi, yaitu fungsi *step* (*step function*) dan fungsi *pulse* (*pulse function*). *Step function* adalah suatu bentuk intervensi yang terjadinya dalam kurun waktu yang panjang. Secara matematik, bentuk intervensi *step function* ini biasanya dinotasikan sebagai berikut:

$$S_t^{(T)} = \begin{cases} 0, & t < T, \\ 1, & t \geq T. \end{cases}$$

dimana T adalah waktu mulainya terjadi intervensi.

Sedangkan *pulse function* adalah suatu bentuk intervensi yang terjadinya hanya dalam suatu waktu tertentu. Secara matematik, bentuk intervensi *pulse function* ini biasanya dinotasikan sebagai berikut:

$$P_t^{(T)} = \begin{cases} 0, & t \neq T, \\ 1, & t = T. \end{cases}$$

dimana T adalah waktu terjadinya intervensi.

2.3 Deteksi Outlier

Outlier adalah data pengamatan yang tidak konsisten pada seriesnya. Efek kejadian tersebut dapat dihitung dengan model intervensi jika waktu dan penyebab diketahui. Ada empat macam jenis outlier yaitu *Innovational Outlier* (IO), *Additive Outlier* (AO), *Temporary Change* (TC), dan *Level Shift* (LS). Deteksi outlier pertama kali dikemukakan Fox (1972) dalam Wei (2006) yang memperkenalkan outlier tipe 1 atau *additive outliers* (AO) dan tipe 2 atau *innovation outliers* (IO).

Additive outlier adalah kejadian yang mempunyai efek pada data time series hanya pada satu periode saja. Bentuk umum sebuah *Additive Outliers* (AO) dalam proses ARMA diuraikan sebagai berikut

$$\begin{aligned} Z_t &= \begin{cases} X_t & t \neq T \\ X_t + \omega & t = T \end{cases} \\ &= X_t + \omega I_t^{(T)} \\ &= \frac{\theta(B)}{\phi(B)} a_t + \omega I_t^{(T)} \end{aligned}$$

dimana

$$I_t^{(T)} = \begin{cases} 1 & t = T \\ 0 & t \neq T \end{cases}$$

adalah variabel indikator yang mewakili ada atau tidak adanya outlier pada waktu T .

Innovational outliers adalah kejadian yang efeknya mengikuti proses ARMA. Bentuk umum sebuah *innovational outliers* didefinisikan sebagai berikut

$$Z_t = X_t + \frac{\theta(B)}{\phi(B)} \omega I_t^{(T)} = \frac{\theta(B)}{\phi(B)} (a_t + \omega I_t^{(T)})$$

Dari persamaan-persamaan tersebut, dapat disimpulkan bahwa *additive outlier* (AO) hanya mempengaruhi pengamatan ke- T , sedangkan *innovational outlier* (IO) mempengaruhi semua pengamatan Z_T, Z_{T+1}, \dots , melebihi waktu T sepanjang memori sistem yang dijelaskan oleh $\theta(B)/\phi(B)$.

Secara umum, sebuah data *time series* bisa saja mengandung beberapa *outlier*, misalnya k buah *outlier* dengan tipe yang berbeda. Sehingga model umum *outlier* dapat ditulis sebagai berikut

$$Z_t = \sum_{j=1}^k \omega_j v_j(B) I_t^{(T_j)} + X_t$$

dimana $X_t = (\theta(B)/\phi(B))a_t$, $v_j(B) = 1$ untuk AO, dan $v_j(B) = \theta(B)/\phi(B)$ untuk IO pada waktu $t = T_j$. (Wei, 2006)

Selain dua tipe outlier tersebut, masih ada dua tipe outlier lain yang sering dibahas dalam analisis time series, yaitu *Level Shift* (LS) dan *Temporary Change* (TC). Suatu LS adalah kejadian yang mempengaruhi deret pada satu waktu tertentu yang memberikan suatu perubahan tiba-tiba dan permanen. Model outlier LS dinyatakan sebagai:

$$Z_t = X_t + \frac{1}{(1-B)} \omega I_t^{(T)}$$

Sedangkan TC adalah suatu kejadian dimana outlier menghasilkan efek awal sebesar ω pada waktu t , kemudian secara perlahan sesuai dengan besarnya δ . Model TC dapat dituliskan sebagai berikut:

$$Z_t = X_t + \frac{1}{(1-\delta B)} \omega I_t^{(T)}$$

Pada saat $\delta = 0$ maka TC akan menjadi kasus additive outlier, sedangkan pada saat $\delta = 1$ maka TC akan menjadi kasus *level shift*.

3. METODOLOGI PENELITIAN

Data yang akan digunakan adalah data skunder yang diambil dari Badan Pusat Statistik tentang data kunjungan wisatawan domestik ke Yogyakarta pada tahun 2006-2010 yang menginap di hotel berbintang dan hotel melati yang ada di seluruh Yogyakarta.

Dugaan awal dari variabel intervensi adalah fungsi *pulse* (*pulse function*) karena pengaruh intervensi hanya terjadi pada periode tertentu. Variabel intervensi yang digunakan terdiri dari satu variabel yaitu Jumlah wisatawan ke Yogyakarta sebelum dan setelah terjadi penurunan drastis.

$$X_t = P_t^{(T)} = \begin{cases} 0, & t \neq T(\text{saat tidak terjadi penurunan drastis wisatawan}) \\ 1, & t = T(\text{saat terjadi penurunan drastis wisatawan}) \end{cases}$$

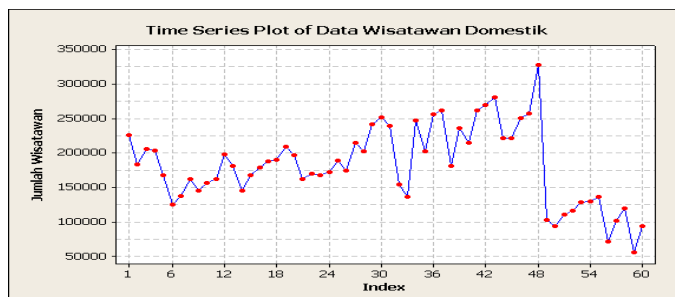
T adalah waktu terjadinya intervensi yaitu terjadinya intervensi turunnya jumlah wisatawan domestik yang datang ke Yogyakarta (tepatnya T mewakili bulan terjadinya penurunan jumlah wisatawan).

Langkah-langkah yang akan dilakukan dalam menganalisis data adalah:

1. Membuat plot *time series* dan kemudian menduga variabel intervensi yang mungkin,
2. Membagi data wisatawan menjadi dua, yaitu data sebelum dan setelah terjadinya intervensi.
3. Membentuk model ARIMA untuk data sebelum terjadinya intervensi.
4. Setelah model ARIMA untuk data sebelum intervensi terbentuk, kemudian dengan model ARIMA tersebut dilakukan peramalan untuk data sebelum intervensi sampai dengan data setelah intervensi terpenuhi dan dihitung residual respon antara data setelah intervensi dengan hasil peramalan dari data sebelum intervensi.
5. Pembentukan model intervensi berdasarkan *plot* residual.
6. Mengestimasi parameter intervensi dan dilakukan pengujian, apabila parameter intervensi tidak signifikan maka model yang digunakan adalah model ARIMA.
7. Setelah model terbentuk selanjutnya dilakukan deteksi outlier.
8. Outlier yang terdeteksi kemudian dimasukkan kedalam model dan dilakukan estimasi parameter.
9. Setelah hasil estimasi parameter model didapatkan selanjutnya dilakukan verifikasi model dengan penambahan outlier. Jika model belum sesuai maka kembali dideteksi outlier, tapi jika model telah sesuai maka bisa digunakan untuk peramalan.
10. Meramalkan data dengan model intervensi (jika signifikan) dan hasil deteksi outlier yang memiliki MSE terkecil.

4. HASIL DAN PEMBAHASAN

Langkah pertama untuk mendapatkan model peramalan indeks harga saham kospi adalah membuat plot *time series*. Hasil plot *time series* dapat dilihat pada Gambar 1.



Gambar 1. Plot Time Series Data Wisatawan Domestik
Periode Januari 2006 – Desember 2010

Berdasarkan gambar 1 tersebut terlihat bahwa sejak Januari 2006 sampai Desember 2009 jumlah kunjungan wisatawan domestik cenderung stabil bahkan mengalami peningkatan jumlah wisatawan domestik. Jumlah kunjungan wisatawan tertinggi terjadi pada bulan ke-48 atau bulan Desember 2009. Namun pada Januari 2010 terjadi penurunan drastis jumlah wisatawan domestik dan setelahnya cenderung stabil. Hal ini mengindikasikan bahwa dalam data wisatawan domestik ke DI Yogyakarta mengandung outlier (pencilan) atau dapat juga mengandung suatu intervensi. Oleh karena itu, untuk data wisatawan domestik ke Di Yogyakarta periode Januari 2006 – Desember 2010 akan dianalisis menggunakan deteksi outlier dan analisis intervensi.

4.1 Pemodelan ARIMA Data Wisatawan Domestik

Pemodelan ARIMA data wisatawan domestik dengan transformasi $\lambda = 0,50$ yaitu transformasi akar kuadrat. model terbaik yang dipilih adalah ARIMA (1,0,[20]) karena memiliki nilai AIC, SBC dan MSE yang lebih kecil dibandingkan dengan nilai AIC, SBC dan MSE model ARIMA lainnya seperti terlihat pada tabel 1

Tabel 1. Pemodelan Terbaik Data Wisatawan Domestik

Model	Signifikansi Parameter	White Noise	Distribusi Normal	Homoskedastisitas	AIC	SBC	MSE
ARIMA(1,0,0)	√	√	x	√	644,0107	648,1994	2512
ARIMA(0,0,3)	√	√	x	√	645,3442	653,7216	2403
ARIMA(1,0,[7])	√	√	x	√	641,4121	647,6951	2327
ARIMA(1,0[13])	√	√	√	√	638,1961	644,4791	2206
ARIMA(1,0,[20])	√	√	√	√	634,2778	640,5609	2066

Sehingga model ARIMA yang terbentuk adalah sebagai berikut

$$\begin{aligned}
 (1 - \phi_1 B)Z_t &= (1 - \theta_{20})a_t \\
 (1 - 0,64165B)Z_t &= (1 - 0,71242B^{20})a_t \\
 Z_t &= 0,64165Z_{t-1} + a_t - 0,71242a_{t-20}
 \end{aligned}$$

4.2 Deteksi Outlier Data Wisatawan Domestik

Untuk mengetahui ada atau tidaknya outlier dalam data, maka dilakukan deteksi outlier. Hasilnya terdeteksi 13 outlier pada data wisatawan, yaitu pada data ke-49, 50, 51, 33, 52, 56, 59, 38, 32, 48, 44, 6 dan 35 yang bertipe Additive Outlier (AO) dan Level Shift (LS). Selanjutnya dilakukan estimasi untuk model ARIMA (1,0,[20]) dengan menambahkan outlier tersebut satu persatu hingga diperoleh model yang memenuhi signifikansi parameter dan juga asumsi residual model, *white noise*, normalitas residual, dan homoskedastisitas.

Model terbaik yang diperoleh adalah ARIMA (1,0,[20]) dengan penambahan delapan outlier. Model tersebut terpilih sebagai model terbaik karena memenuhi semua uji yang dilakukan yaitu signifikansi parameter, uji *white noise* residual, normal residual, dan homoskedastisitas residual, meskipun belum semua outlier ditambahkan kedalam model. Hasil estimasinya dapat dilihat pada tabel 2 sebagai berikut.

Tabel 2. Pemodelan Terbaik Data Wisatawan Domestik

Parameter	Type	Estimate	Std. Error	t Value	P value
μ	-	436,851	14,7949	29,53	<.0001
θ_{20}	-	0,53701	0,19366	2,77	0,0078
ψ_1	-	0,83261	0,09887	8,42	<.0001
I_{49}	Level Shift	-191,053	34,65575	-5,51	<.0001
I_{50}	Level Shift	-155,699	43,25446	-3,6	0,0007
I_{51}	Level Shift	-93,0383	43,07772	-2,16	0,0357
I_{33}	Additive	-84,4889	24,22899	-3,49	0,001
I_{52}	Level Shift	-84,7983	35,79337	-2,37	0,0218
I_{56}	Level Shift	-63,2797	26,71409	-2,37	0,0218
I_{59}	Additive	-82,8283	27,3811	-3,03	0,004
I_{38}	Additive	-65,837	23,78751	-2,77	0,0079

Jadi model yang terbentuk untuk data wisatawan domestik adalah sebagai berikut.

$$Z_t = 0,64165Z_{t-1} + a_t - 0,71242a_{t-20} - 191,053I_T^{(49)} - 155,699I_T^{(50)} - 93,0383I_T^{(51)} - 84,4889I_T^{(33)} - 84,7983I_T^{(52)} - 63,2797I_T^{(56)} - 82,8283I_T^{(59)} - 65,837I_T^{(38)}$$

4.3 Analisis Intervensi Data Wisatawan Domestik

Pemodelan pre intervensi data wisatawan domestik dilakukan pada data ke-1 sampai ke ke-48 dengan diferensi dan transformasi $\lambda=0,50$. Model pre intervensi terbaik adalah ARIMA ([12],1,0) dengan MSE sebesar 1325. Secara lengkap dapat dilihat pada tabel 3 sebagai berikut.

Tabel 3. Pemodelan Terbaik sebelum Intervensi

Model	Signifikansi Parameter	White noise	Distribusi Normal	Homoskedastisitas	AIC	SBC	MSE
ARIMA([12],1,0)	√	√	√	x	473,282	475,132	1325
ARIMA(0,1,1)	√	√	√	√	477,998	479,848	1465
ARIMA(0,1,[10])	√	x	√	√	479,673	481,523	1518
ARIMA(0,1,[12])	√	√	√	x	474,568	476,418	1362
ARIMA(0,1,[10,12])	√	√	√	x	472,766	476,466	1256

Tabel 4 menunjukkan estimasi parameter dan uji signifikansi pemodelan intervensi. Pada pemodelan ini, orde intervensi yang didapatkan berdasarkan plot residual, terbentuk tiga fungsi transfer sebagai berikut fungsi transfer pertama yaitu $b=0, s=0, r=1$, fungsi transfer kedua yaitu $b=7, s=0, r=0$ serta fungsi transfer ketiga yaitu $b=10, s=0, r=0$. Model intervensi ini memenuhi semua uji yang dilakukan yaitu signifikansi parameter, uji *white noise* residual, normal residual, dan homoskedastisitas residual.

Tabel 4. Estimasi Parameter Model Intervensi

Model	Parameter	Estimate	Std. Error	t Value	P value
ARIMA ([12],1,0) $b=0, s=0, r=1; b=7, s=0, r=0; b=10, s=0, r=0$	ψ_{12}	0,47479	0,14363	3,31	0,0019
	ω_0	-254,19	36,134	-7,03	<.0001
	δ_1	1,0201	0,03493	29,2	<.0001
	ω_7	-61,776	25,9145	-2,38	0,0215

Jadi model intervensi yang terbentuk untuk data wisatawan domestik adalah sebagai berikut.

$$Z_t = \frac{(\omega_0 B^0)}{1 - \delta_1 B} P_t^{(T)} + (\omega_0 B^7) P_t^{(T)} + (\omega_0 B^{10}) P_t^{(T)} + \frac{1}{(1 - B)(1 - \phi_{12} B^{12})} a_t$$

$$Z_t = \frac{(-254,19 B^0)}{(1 - 1,0201 B)} P_t^{(T)} + (-61,776 B^7) P_t^{(T)} + (-77,077 B^{10}) P_t^{(T)} + \frac{1}{(1 - B)(1 - 0,47479 B^{12})} a_t$$

4.4 Analisis Intervensi dan Deteksi Outlier

Setelah terbentuk model intervensi untuk data wisatawan domestik ke DI Yogyakarta, selanjutnya untuk mengetahui ada atau tidaknya outlier dalam data, maka dilakukan deteksi outlier. Dari tabel 35 terlihat bahwa terdeteksi lima outlier pada data wisatawan, yaitu pada data ke-34, 38, 32, 33 dan 24 yang bertipe Additive Outlier (AO) dan Level Shift (LS). Selanjutnya dilakukan estimasi untuk model ARIMA ([12],1,0) dengan menambahkan outlier tersebut satu persatu hingga diperoleh model yang memenuhi signifikansi parameter dan juga asumsi residual model, *white noise*, normalitas residual, dan homoskedastisitas.

Model terbaik yang diperoleh adalah model intervensi ARIMA ([12],1,0) dengan fungsi transfer pertama yaitu $b=0, s=0, r=1$, fungsi transfer kedua yaitu $b=7, s=0, r=0$ serta fungsi transfer ketiga yaitu $b=10, s=0, r=0$ serta penambahan empat outlier. Model tersebut terpilih sebagai model terbaik karena memenuhi semua uji yang dilakukan yaitu signifikansi parameter, uji *white noise* residual, normal residual, dan homoskedastisitas residual, meskipun belum semua outlier ditambahkan kedalam model. Estimasi parameter model terdapat pada tabel 5.

Tabel 5. Estimasi Parameter Model ARIMA dengan Delapan Outlier

Parameter	Type	Estimate	Std. Error	t Value	P value
ψ_{12}	-	0,5976	0,1579	3,78	0,0005
ω_0	-	-254,76	25,342	-10,05	<.0001
δ_1	-	1,02356	0,02374	43,11	<.0001
ω_7	-	-58,059	18,5387	-3,13	0,0032
ω_{10}	-	-73,528	18,4844	-3,98	0,0003
I_{34}	Additive	47,6399	18,8357	2,53	0,0155
I_{38}	Level Shift	-49,534	15,5138	-3,19	0,0027
I_{32}	Level Shift	-62,442	18,8949	-3,3	0,002
I_{33}	Additive	-66,442	21,7751	-3,05	0,004

Jadi model yang terbentuk untuk data wisatawan domestik adalah sebagai berikut.

$$Z_t = \frac{(-254,76B^0)}{(1 - 1,02356B)}P_t^{(T)} + (-58,059B^7)P_t^{(T)} + (-73,528B^{10})P_t^{(T)} + \frac{1}{(1 - B)(1 - 0,5976B^{12})}a_t$$

$$+ 47,6399I_T^{(34)} - 49,534I_T^{(38)} - 62,442I_T^{(32)} - 66,442I_T^{(33)}$$

4.5 Perbandingan Model Outlier dengan Model Intervensi

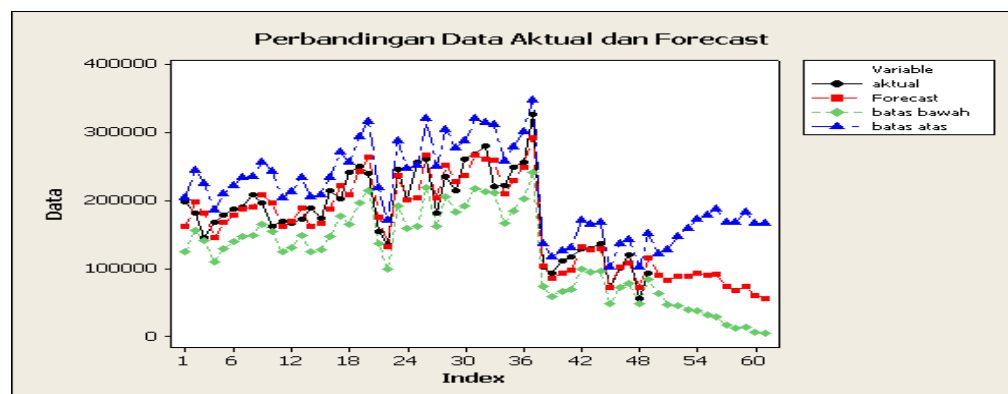
Setelah semua analisis yang dilakukan untuk data wisatawan domestik ke DI Yogyakarta, pada tahap ini akan dilakukan perbandingan nilai MSE dari semua model yang didapat dari hasil analisis sebelumnya. Berdasarkan tabel 6 terlihat bahwa MSE terkecil yang dihasilkan terdapat pada model ARIMA dengan intervensi dan outlier, yaitu ARIMA ([12],1,0) b=0, s=0, r=1; b=7, s=0 r=0; b=10, s=0 r=1 dengan empat outlier. Sehingga model tersebut dikatakan model terbaik dan digunakan untuk peramalan.

Tabel 6. Perbandingan Nilai MSE Model Intervensi dan Outlier

Model	MSE
ARIMA(1,0,[20])	2066
ARIMA(1,0,[20]) dengan delapan outlier	948,0842
ARIMA ([12],1,0) b=0, s=0, r=1; b=7, s=0 r=0; b=10, s=0 r=0	1172
ARIMA ([12],1,0) b=0, s=0, r=1; b=7, s=0 r=0; b=10, s=0 r=1 dengan empat outlier	523,7167

4.6 Peramalan Data Wisatawan Domestik

Hasil ramalan jumlah kunjungan wisatawan berada di dalam ambang batas (95% *confident interval*). Sehingga hasil ramalan dari model intervensi fungsi *pulsed* dengan penambahan empat outlier dapat digunakan untuk memperkirakan jumlah wisatawan domestik pada bulan Januari - Desember 2011 pasca intervensi penurunan drastis jumlah wisatawan pada bulan Januari 2010. Pada gambar 2 juga diperlihatkan grafik perbandingan antara nilai aktual dengan nilai peramalan dari data wisatawan domestik ke DI Yogyakarta menggunakan model intervensi dan deteksi outlier. Pada gambar tersebut terlihat bahwa nilai peramalannya mendekati nilai aktual data, sehingga model intervensi dan deteksi outlier yang digunakan dapat dikatakan sudah baik.



Gambar 2. Perbandingan Nilai Aktual dan Forecast Data Wisatawan Domestik

5. KESIMPULAN

5.1 Kesimpulan

Berdasarkan analisis dan pembahasan yang telah dilakukan, maka dapat disimpulkan bahwa model terbaik yang dipilih adalah model intervensi dengan penambahan empat outlier dengan model sebagai berikut.

$$Z_t = \left(\frac{(-254,76B^0)}{(1 - 1,02356B)} - 58,059B^7 - 73,528B^{10} \right) P_t^{(T)} + \frac{1}{(1 - B)(1 - 0,5976B^{12})} a_t \\ + 47,6399I_T^{(34)} - 49,534I_T^{(38)} - 62,442I_T^{(32)} - 66,442I_T^{(33)}$$

Peramalan yang dilakukan merupakan peramalan bulanan untuk 12 bulan ke depan. Hasil dari *forecasting* tersebut dapat digunakan karena terletak dalam selang kepercayaan 95%. Selain itu, nilai peramalannya juga mendekati nilai aktual data, sehingga model intervensi dan deteksi outlier yang digunakan dapat dikatakan sudah baik.

5.2 Saran

Karena model sebelum intervensi untuk data wisatawan domestik yang dibahas pada tugas akhir ini yaitu ARIMA ([12],1,0) belum memenuhi asumsi homoskedastisitas residual, maka untuk peneliti selanjutnya dapat dimodelkan bukan hanya pemodelan mean tapi juga pemodelan terhadap variannya dengan menggunakan ARCH atau GARCH untuk mengatasi heteroskedastisitas yang terjadi dalam residual model.

DAFTAR PUSTAKA

- Box, G.E.P., Jenkins, G.M., and Reissel.G.C. 1994. *Time Series Analysis Forecasting and Control*. 3rd edition. Englewood Cliffs: Prentice Hall.
- Box, G.E.P and Tiao, G.C. 1975. *Intervention Analysis with Applications to Economic and Environmental Problems*. Journal of American Statistics Association, 70, pp. 70-79.
- Daniel, W.W. (1989), *Statistika Nonparametrik Terapan*, PT. Gramedia, Jakarta.
- Hamilton, J.D. 1994. *Time Series Analysis*. New Jersey. Princeton University Press.
- Makridakis, S., Wheelwright, S.C., and McGee, V.E. 1999. *Jilid 1 edisi kedua, Terjemahan Ir. Hari Suminto, Metode dan Aplikasi Peramalan*. Jakarta. Bina Rupa Aksara.
- Soejoeti, Z. 1987. *Analisis Runtun Waktu*, Materi Pokok UT. Jakarta. Karunika.
- Suhartono. 2007. *Teori dan Aplikasi Model Intervensi Fungsi Pulse*. Jurnal Ilmiah MatStat, Vol.7, No.2
- Suhartono dan Wahyuni, W. 2002. *Analisis Dampak Promosi dan Kenaikan Harga terhadap Fluktuasi Jumlah Pelanggan dan Pemakaian Pulsa di PT. Telkom Divre V*. Forum Statistika dan Komputasi. Edisi Khusus Seminar Nasional Statistika, IPB: Bogor.
- Tsay, RS. 2002. *Analysis of Financial Time Series*. John Wiley and Sons, Inc. Chicago.
- Wei, W.W.S. 2006. *Time Series Analysis, Univariate and Multivariate Methods*. Canada. Addison Wesley Publishing Company.